Time and frequency domain characteristics of sperm whale clicks

John C. Goold, and Sarah E. Jones

Citation: The Journal of the Acoustical Society of America **98**, 1279 (1995); doi: 10.1121/1.413465
View online: https://doi.org/10.1121/1.413465
View Table of Contents: https://asa.scitation.org/toc/jas/98/3
Published by the Acoustical Society of America

Time and frequency domain characteristics of sperm whale clicks

John C. Goold and Sarah E. Jones
University College of North Wales, School of Ocean Sciences, Menai Bridge, Bangor, Gwynedd LL59 5EY, United Kingdom

(Received 17 October 1994; accepted for publication 11 April 1995)

Regular clicks from diving sperm whales, both large bull males and smaller females, were recorded in deep oceanic water off the Azores and subsequently sampled to computer disks for digital analysis. A total of 8540 clicks were marked and analyzed. Simple temporal analysis of the interclick intervals during feeding dives revealed mean click rates for male sperm whales of 1.1713 s^{-1} and 1.9455 s^{-1} for females. Fourier analysis showed distinctive peaks in the spectra of bull male sperm whales at 400 Hz and 2 kHz which were stable over extended periods of up to 20 mins. The clicks contained higher frequency components with energy ranging up to at least 12 kHz but not concentrated at any sharply defined frequency. The clicks of smaller female sperm whales showed similar spectral peaks, shifted to 1.2 and 3 kHz, respectively, but these peaks were less pronounced than those in the male click spectra and less stable with time. Higher frequencies were also present up to at least 15 kHz. The previously reported multiple pulse structure of sperm whale clicks is confirmed, but digital filtering reveals this structure to be frequency dependent. Analysis using the short-time Fourier transform confirms the complex time–frequency structure of individual clicks. The frequencies at which the multiples emerge in male and female clicks supports the idea of air cavities in the sperm whale head acting as sound reflectors, although the magnitude of the second pulse at high frequencies suggests some form of off axis distortion. It is also possible that air cavity resonance in the head of the sperm whale may act to reinforce the high-frequency components of the click, and that such components may have superior range and resolution performance in terms of echolocation.

PACS numbers: 43.80.Ka, 43.80.Lb

LIST OF SYMBOLS

n	positive integer
DT	detection threshold (dB)
SL	source level (dB)
TL	transmission loss (dB)
TS	target strength (dB)
NL	isotropic noise spectral density (dB)
DI	receive beam directivity index (dB)
r	range (meters)
V	volts
s	seconds
c	sound velocity
FFT	fast Fourier transform
STFT	short-time Fourier transform
FIR	finite impulse response
IPI	interpulse interval

INTRODUCTION

During the course of the last four decades acoustic emissions from sperm whales (*Physeter macrocephalus L.*) have generated much interest. Sperm whale sounds were first recorded in the 1950s and later identified (Worthington and Schevill, 1957). These workers reported that sperm whales produce loud, impulsive "clicks" at various repetition rates, an observation that has since been confirmed by several workers (Backus and Schevill, 1966; Bunsel and Dziedzic, 1967; Norris and Harvey, 1972; Gordon, 1987; Weilgart and Whitehead, 1988). A variety of sounds have been attributed to sperm whales (see Gordon, 1987 and Weilgart, 1990 for a review): the present study, however, concentrates only on the so-called "regular clicks." Sperm whales are prodigious divers and frequently undertake vertical excursions in excess of 500 m in search of food, primarily deep sea squid (Clarke *et al.*, 1993). Dives of even greater vertical extent, possibly in excess of 2000 m, have also been reported (Lockeyer, 1977; Watkins *et al.*, 1993). While undertaking these dives, which may last up to 90 min, animals invariably produce a monotonous series of clicks with repetition rates between 0.5 and 2 clicks per second. It is generally assumed that the primary function of these sounds is echolocation, and at intervals during the dive click rates increase markedly, leading to very "faint" but rapid bursts of clicks with repetition rates of up to 200 s^{-1}. These rapid bursts are commonly referred to as "creaks," and the use of this term will be continued here. The regular clicks are broadband transient signals, with a duration of some 10–20 ms depending upon the size of the animal producing them. Watkins (1977) noted click energy ranging from below 100 Hz to well in excess of 20 kHz, with major emphasis in the 2 to 6 kHz range. Other workers report dominant frequencies of 5 kHz (Backus and Schevill, 1966), 1 kHz (Bunsel and Dziedzic, 1967), 2–8 kHz (Levenson, 1974), and 2 kHz (Weilgart and Whitehead, 1988). Levenson (1974) calculated a mean broadband (250 Hz–16 kHz) source level for sperm whale clicks at 171.2 dB *re*: 1

 0001-4966/95/98(3)/1279/13/$6.00

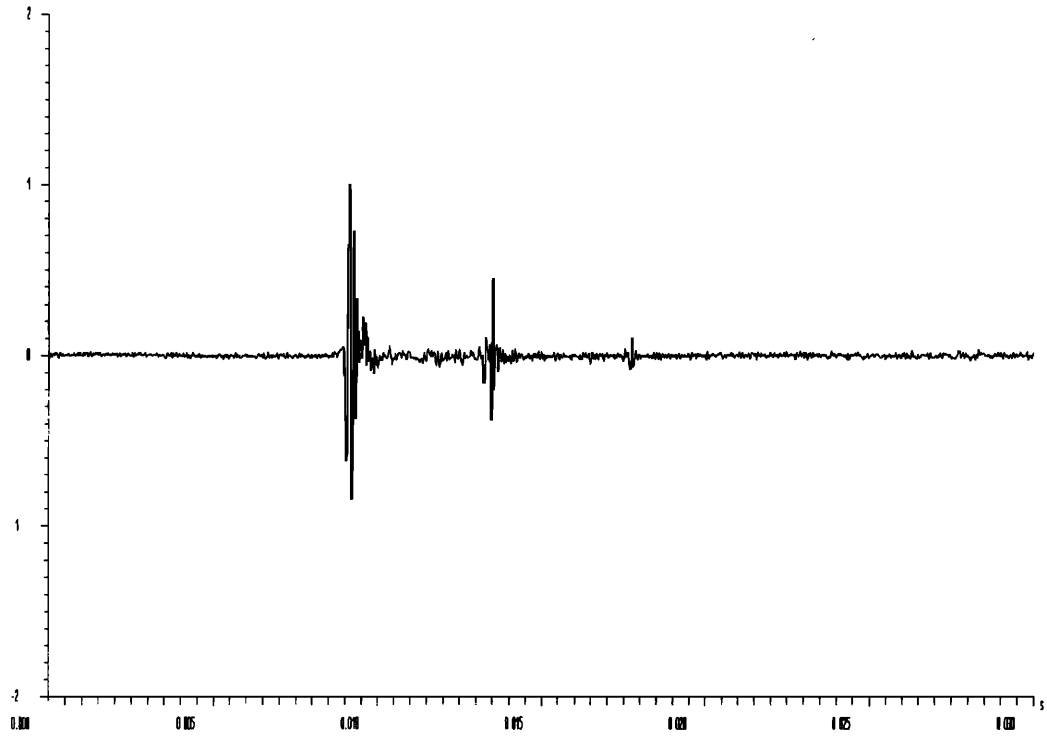

FIG. 1. Example click from a diving female sperm whale showing regularly spaced pulses with decaying amplitude. Time base is 30 ms; waveform units in volts.

μPa, although it is not stated whether this refers to peak or rms values. Individual clicks within creaks have received little attention in the literature, and their presence is merely noted in this study.

One striking feature of sperm whale clicks, noted by Backus and Schevill (1966), is that they consist of a series of regularly spaced pulses with decaying amplitude. An example of such a click, recorded from a diving female sperm whale, is shown in Fig. 1. In a classic paper Norris and Harvey (1972) attributed this pulsed structure to multiple reflection of an initial sound pulse from air sacs within the head of the sperm whale. The sperm whale head is largely occupied by the spermaceti organ, a complex mass of oil filled connective tissue, enclosed by a membrane called the case, and surrounded by layers of muscle and blubber (Fig. 2). The two major components of this structure are the spermaceti sac, which contains tissue rich in spermaceti oil, and the junk, which is composed of alternating segments of spermaceti tissue and tough connective tissue, the latter probably serving a structural role. The left nasal passage runs from the blow hole down the left side of the spermaceti organ, and down through a hole in the base of the skull, thus forming a direct pathway for respiration. The right nasal passage runs as a broad, flattened tube through the tissues of the spermaceti organ, opening into two air cavities, the nasofrontal and distal sacs, at the front and rear of the spermaceti organ. These air cavities are proposed to act as sound reflectors due to the large difference in acoustic impedance between air and spermaceti oil, even at the elevated pressures encountered at depths of 2000 m. Norris and Harvey hypothesized that sound is produced at the front of the sperm whale head by pneumatic action on a structure known as the museau de singe (monkey's muzzle). The museau de singe is a tough, fibrous structure in the form of two tightly occluding lips at the anterior end of the right nasal passage, where it opens out into the distal sac. Air is, supposedly, forced forward along the broad flattened tube of the right nasal passage and through the lips of the museau de singe, thereby producing a loud, impulsive burst of sound. As the left and right nasal passages are cross-connected via narrow airways near the blow hole and at the base of the frontal sac, air used in sound production could be recycled to allow continuous click production with a finite quantity of air. As the nasofrontal and distal sacs bound the ends of the long, oil-filled spermaceti sac along the whales' longitudinal axis, it was suggested that

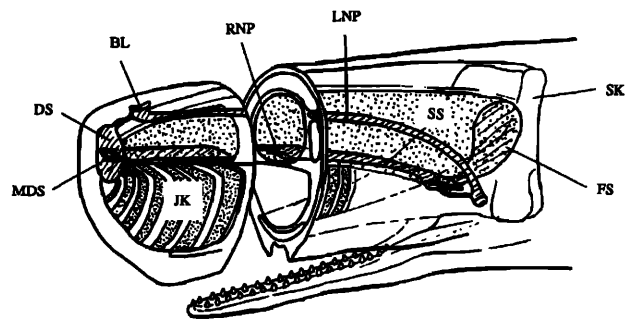

FIG. 2. Anatomy of the sperm whale head, drawn after Clarke (1979) copyright (©) 1979 by Scientific American Inc., all rights reserved. MDS—museau de singe, DS—distal sac, BL—blowhole, RNP—right nasal passage, LNP—left nasal passage, SK—skull, FS—frontal sac, SS—spermaceti sac, JK—junk.

the pulsed structure of the clicks results from longitudinal reverberation of the initial sound pulse partially reflected from these air cavities. Therefore, it follows that the time delay between successive pulses within a click should be a function of distance separating the frontal and distal air sacs, and the velocity of sound in spermaceti oil. In other words, the time delay between the pulses is directly proportional to the length of the spermaceti organ, and therefore also a function of the total body length of the animal. A few studies since have suggested this to be the case (Mohl *et al.*, 1981; Alder-Fenchel, 1980; Gordon, 1991; Mohl and Amundin, 1991), and an empirical relationship to relate the so-called interpulse interval (IPI) to the spermaceti sac length, and hence body length, of individual animals is quoted in Gordon (1991). Since Norris and Harvey's paper, several workers have studied sperm whales in their deep ocean environment, often employing hydrophones and recording equipment for the purposes of tracking and acoustic studies (Watkins, 1977; Gordon, 1987a; Weilgart and Whitehead, 1988; Whitehead and Weilgart, 1990). Patterns of clicking have been investigated and the clicks themselves have been subject to spectral analysis. Although the pulsed structure is frequently commented upon, little detailed analysis of the clicks themselves has been performed, and the sound production mechanism still remains a hypothesis. The present work uses high-quality recordings of sperm whale clicks to investigate the detailed time and frequency domain characteristics of clicks from both large bull male sperm whales and the smaller females.

I. METHODS

A. Field data collection

From 1987–1993 The International Fund for Animal Welfare (IFAW) sponsored benign research into the behavior and ecology of sperm whales in the waters around the Azores, a group of nine North Atlantic islands situated between latitudes 36° 50′N–39° 50′N and longitudes 31°30′W–24°50′W. Regular clicks of diving sperm whales were recorded from the IFAW research vessel SONG OF THE WHALE, a 14-m Ketch, through the field seasons 1987–1993. Stereo recordings were made on magnetic tape through towed hydrophones. The hydrophone assembly consisted of two Benthos AQ-4 transducers, each with separate preamplifiers and integral 200-Hz high-pass filters, mounted 20 m apart along the axis of the tow line in a plastic tube filled with castor oil for close impedance matching to seawater. The assembly was towed on 100 m of strengthened cable behind the vessel. When under sail, the array geometry was simply two elements 100 and 120 m aft; when stationary, the hydrophone assembly and tow cable sank thereby resulting in a geometry of elements 100 and 120 m directly below the vessel. During the field seasons 1987–1990 the signals were recorded on a Nagra IV reel to reel tape machine; from 1991 onward signals were recorded on digital audio tape (DAT). The frequency response of the AQ-4 elements is flat ±1 dB from 1 Hz up to 15 kHz, and flat ±3 dB from 15 kHz up to 30 kHz. The response of the Nagra is flat ±1 dB from 25 Hz up to 35 kHz at maximum tape speed, although recordings were generally made at half-speed where the response is flat up to 20 kHz. The DAT system is flat ±1 dB from 20 Hz to 22 kHz, and cuts off above 24 kHz. For all clicks analyzed in this report the frequency response of the system can be considered flat from 200 Hz to 15 kHz.

Typically, during daylight encounters, the boat was positioned some 300 m behind a target animal, or as close as possible without causing disturbance to the animal, and maintained this position until the whale "fluked up" and commenced its dive. The action of fluking up usually signifies the beginning of a feeding dive, where the animal brings its tail flukes clear of the water as it assumes a near vertical, head downward posture and descends beneath the surface. A tail fluke photograph was always taken at this point for later identification of individual animals. Once a target animal was submerged, a continuous stereo recording was made of its clicks through the hydrophone array, ideally with the vessel holding station and allowing the hydrophones to descend as far as possible from the surface. Deep hydrophones maximize the time delay between the reception of a click at the elements and reception of a "ghost click" after the sound has been reflected back down onto the elements from the ocean surface, which acts as an almost perfect sound reflector. From 1990 onward a Furuno paper trace depth sounder was installed in the boat and could be used to measure the descent rate of target animals. Immediately after a whale flukes up, a characteristic patch of disturbed surface water, known as the "slick" or "footprint" is visible and conveniently marks the point of the whale's submergence. The vessel could be moved quickly onto the slick, and the echo sounder activated to yield one strong echo from the target whale. The echo returning from the whale was traced on a moving roll of paper, calibrated vertically in fathoms and horizontally in minutes. The depth sounder operated to a maximum depth of 500 fathoms, although it was generally only possible to track the descent of sperm whales over the first half of this range. As whales were generally noted to dive at about 100 m/min, and began clicking some 2 to 3 min after fluking up, all clicks recorded through the hydrophone array were being received from behind the whales during the early dive stages. Unfortunately, the recording system was not calibrated and a transient crew composition often meant that gain settings were not noted, so estimates of click source levels have not been possible, even though the animal's range is known in some instances.

B. Laboratory data acquisition

High-quality recordings of sperm whale clicks were selected from the field data, and copied from the master tapes onto Digital Audio Tape DAT. The selection criteria were (a) single, or very few, whale vocalizations making it possible to follow with confidence the clicks of a particular individual over extended time periods and (b) click sequences with a high signal-to-noise ratio. A total of nine individual click sequences were chosen for analysis and are detailed in Table I. Recordings were sampled onto computer hard disk using a Cambridge Electronic Design (CED) 1401 laboratory interface. The analog line output from the DAT recorder was connected to the sampler input and the signals redigitized for

TABLE I. General information on the nine sperm whale click sequences selected for analysis.

Click sequence code	Whale sex	Sequence duration (min)	Total number of clicks
SW1	♂	20.7	1056
SW2	♂	20.1	1125
SW3	♂	11.0	655
SW4	♂	14.0	775
SW5	♂	13.9	975
SW6	♀	10.8	1032
SW7	♀	11.1	664
SW8	♀	7.4	547
SW9	♀	15.0	1711

storage on disk. A sampling frequency of 62.5 kHz was chosen, thereby comfortably oversampling the maximum frequency reproducible by the DAT medium (24 kHz).

C. Data analysis

Analysis of the sampled waveform data was performed using CED Spike 2 software V.4.70. The features studied were (i) rates of clicking, (ii) frequency domain characteristics of clicks, (iii) time domain characteristics of clicks, and (iv) combined time–frequency characteristics of clicks.

1. Click rates

For each animal the click sequence was viewed sequentially and the onset of each click marked by inspection. Click onset times were then used to calculate the intervals between successive clicks by simple subtraction. Time intervals between clicks of greater that 4 s were taken as intervals between subsequences of clicks, and as such were not included in the calculation of interclick intervals. Reciprocal values were taken to give click rates.

2. Frequency domain analysis

Each click, marked by previous inspection, was transformed to the frequency domain in the form of a power spectrum. Stored click onset times were used as reference points for the analysis of short waveform segments (i.e., individual clicks). The analysis employed a 2048-point fast Fourier transform (FFT) on 32.768 ms of waveform data, beginning at a point approximately 8 ms preceding each click onset, thus encompassing all significant components of the click (Fig. 3 shows typical sperm whales clicks viewed in this way). As the clicks were not significantly truncated, no windowing of waveform data was performed prior to analysis. The FFT routine employed a radix-2 method across 2^n data points, producing $n/2+1$ power magnitude estimates and $n/2-1$ phase estimates. Thus, with a sampling frequency of 62.5 kHz, a 2048-point FFT produced 1025 magnitude estimates, each with a resolution of approximately 30.5 Hz. Magnitude estimates were stored to 11 decimal places and phase estimates were discarded. The power magnitude data, held in ASCII files, were converted to random access format for ease of access and compression of data. Power magnitude estimates were displayed in the form of a 3-D cascade of spectral graphs along a time axis. To display all the spectral data in this manner would require a very large number of plots, so an attempt was made to reduce the bulk of data for display purposes. Each sequence of click spectra was normalized and successive blocks of ten spectra averaged, in order to produce single plots showing an overview of the spectral distribution of energy through the full duration of the respective click sequences.

3. Time domain analysis

The time domain characteristics of the clicks were investigated by selective filtering of the wave data files to pass only the components in selected frequency bands. Digital FIR filters were used, filter coefficients being generated using the Kaiser design parameters, with emphasis placed on sharp passband–stopband transition and low sidelobe ripple levels. Typically filters were generated with passband–stopband transition widths of 3° ($\pi/30$ rad), and maximum sidelobe ripples below −38 dB. Stopband attenuation was always in excess of 50 dB. Once a suitable filter was designed, the coefficient values, in the form of a symmetrical impulse response, were used to filter the waveform data file by digital convolution (Lynn and Fuerst, 1992). As an odd number of filter coefficients was always used (usually $n_{\text{coeff}}=251$), and filtering began $n_{\text{coeff}}/2+1$ points into the wavedata file; the filters exhibited zero phase shift.

4. Time/frequency analysis

Selected sections of clicks were subjected to time/frequency analysis using a short-time Fourier transform in order to investigate changes in the levels and distribution in spectral energy within individual clicks as a function of time. The technique employed a sliding FFT window that was short in comparison to the waveform section of interest (i.e., the whale click). The window was stepped through the wave data sequentially from start to finish, calculating a FFT for each frame. A 256-point (approximately 4 ms) window was used, the right-hand edge beginning at a point immediately preceding the click onset, so as to begin on a "blank" section of wave data. STFT analysis was performed across a total of 18 ms of the whale click, advancing in steps of 0.32 ms. For each step, the 256 wave data points in the analysis window were modulated with a raised cosine function, then transformed to a power spectrum with a 256-point FFT. Compensation was made for a total power loss of 3/8, caused by the cosine function after each FFT. The resulting power magnitude estimates and time shift values were stored and presented in a waterfall display.

II. RESULTS

A. Click rates

Figure 4 shows click rates of the male sperm whales SW1 and SW2 and Fig. 5 shows click rates of the female sperm whales SW6 and SW7; each plot shows click rates as reciprocal values of interclick intervals versus the time since

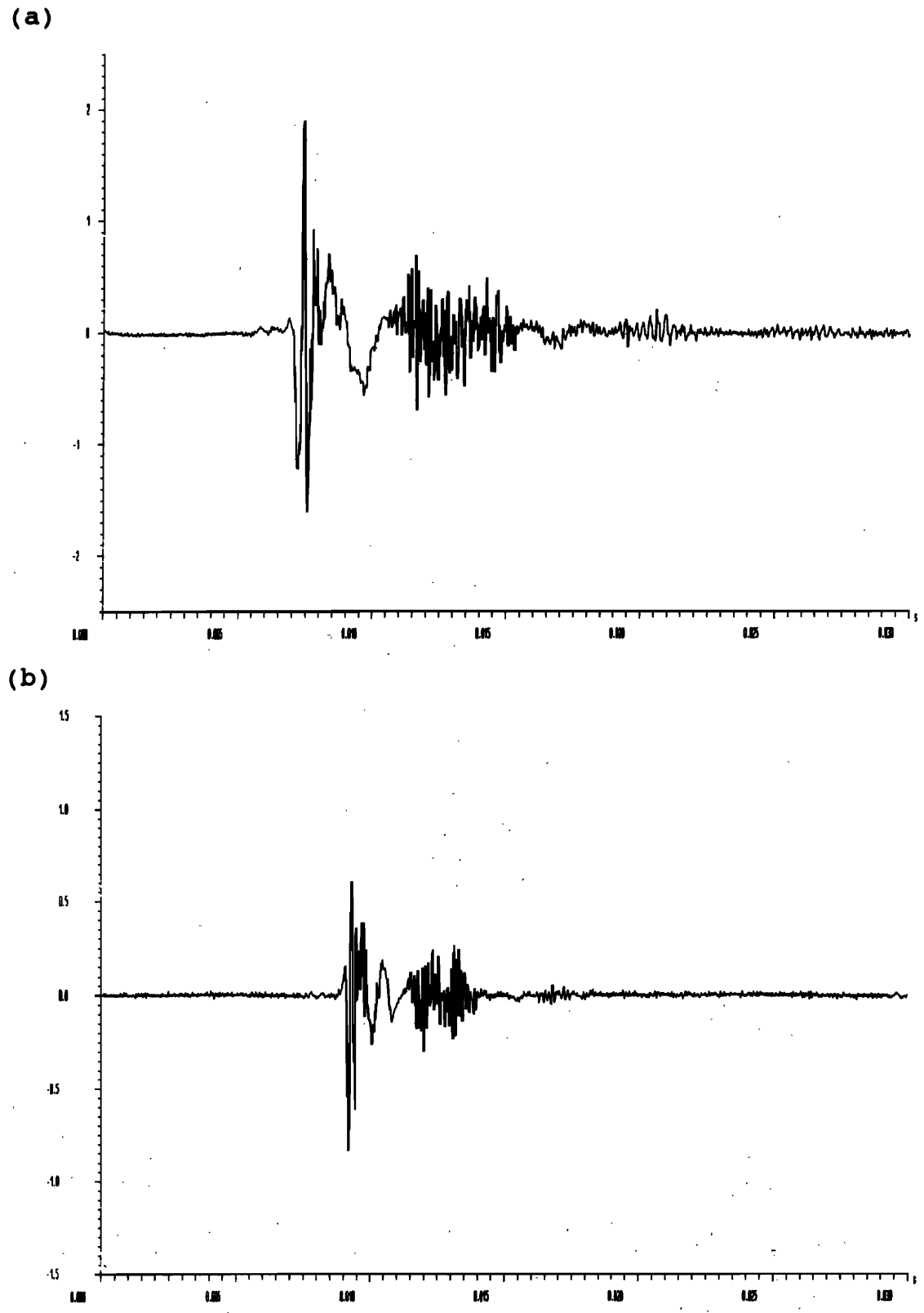

FIG. 3. Typical sperm whale clicks recorded during first few minutes of feeding dives: (a) click from large bull male, (b) click from smaller female. Time base is 30 ms on both plots; waveform units in volts.

the first click. Rates have been smoothed with a ten-point moving average for the purposes of visual clarity. The time lag between a whale fluking up and the first click was usually only some 2 to 3 min, so the time axes could, as a crude approximation, be considered as the time during the dive cycle. The figures show that click rates generally begin between 0.5 and 1 click s^{-1} and increase gradually over the first 2–4 min, whereupon click rates tend to oscillate around an equilibrium level, with the notable exception of SW7 [Fig. 5(b)] where click rate increases over the entire se-

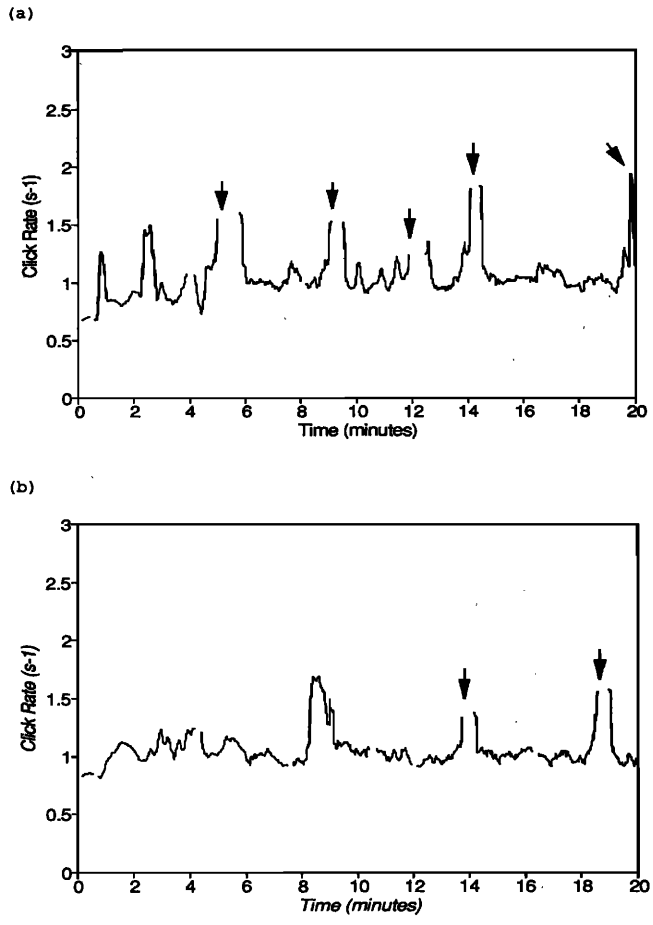

FIG. 4. Click rates from feeding dives of two male sperm whales (a) SW1 and (b) SW2. Data sets plotted using ten-point moving averages for visual clarity and axes scaled identically for comparative purposes. Arrows denote the presence of creaks.

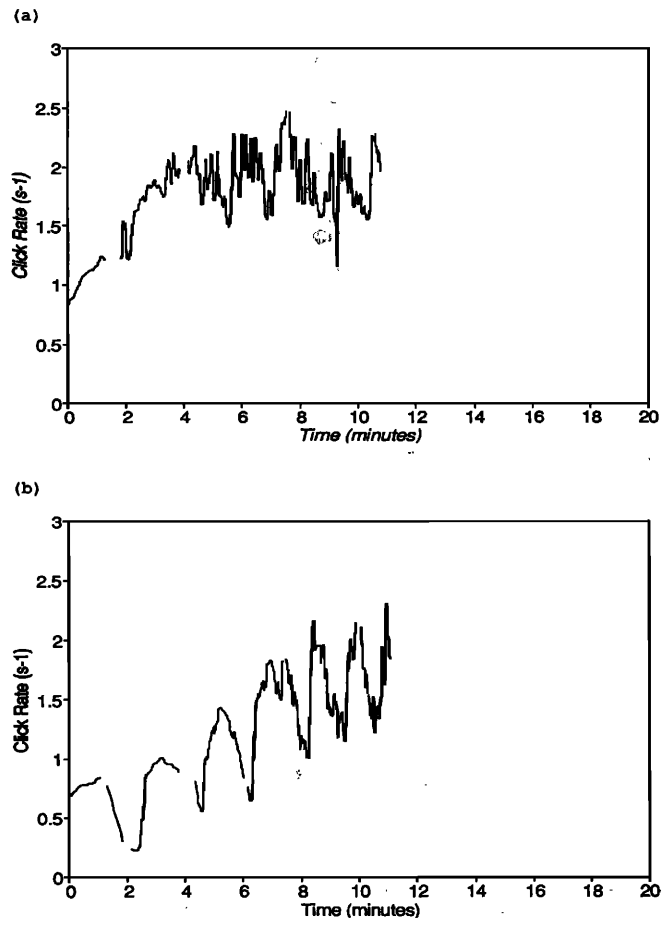

FIG. 5. Click rates from feeding dives of two female sperm whales (a) SW6 and (b) SW7. Data sets plotted using ten-point moving averages for visual clarity and axes scaled identically for comparative purposes.

quence duration of 11 min. Interestingly the amplitudes of oscillations in click rate are greater in the female sequences than in the male sequences. As the graphs are smoothed, very rapid oscillations from one click to the next have been removed, but they are present. During the male sequences sharp increases in click rate can be seen every few minutes; the majority of these immediately precede creaks. If creaks signify final feeding events, as an animal homes in on its prey, such rate increases might be expected.

Mean equilibrium click rates have been calculated for each sequence, using the data from 4 min onward in each case, and the results are shown in Table II. The large bull males SW1–SW4 have the slowest click rates and the females SW6–SW9 the highest. Interestingly SW5 was judged (from its IPI) to be a male smaller than a bull, but larger than the females, and its mean click rate lies between the other two sets. The confidence limits in Table II indicate that the mean click rates of SW1, SW2, and SW4 are all similar at the 95% level, SW3 is significantly greater than all three, and SW5 is greater still. The mean click rates of SW6–SW9 cover a broader range but retain tight confidence limits individually. All the mean click rates of females are significantly different from each other at the 95% level, but are all significantly greater than the males. Pooling the male and female click rate data into two separate samples a one-way analysis of variance shows that male and female click rates are significantly different with a probability value of less than 0.000 and an F value of 8162. Using the pooled data the overall mean click rate for males is 1.1713 s^{-1} (standard error

TABLE II. Mean equilibrium click rates for each sequence, with upper and lower 95% confidence limits.

Click sequence code	Mean equilibrium click rate	Lower confidence limit	Upper confidence limit
SW1	1.0706	1.0554	1.0858
SW2	1.0691	1.0548	1.0835
SW3	1.2943	1.2591	1.3295
SW4	1.0944	1.0790	1.1099
SW5	1.4044	1.3865	1.4223
SW6	1.9205	1.8991	1.9419
SW7	1.5235	1.4860	1.5611
SW8	1.8270	1.8076	1.8464
SW9	2.1543	2.1372	2.1714

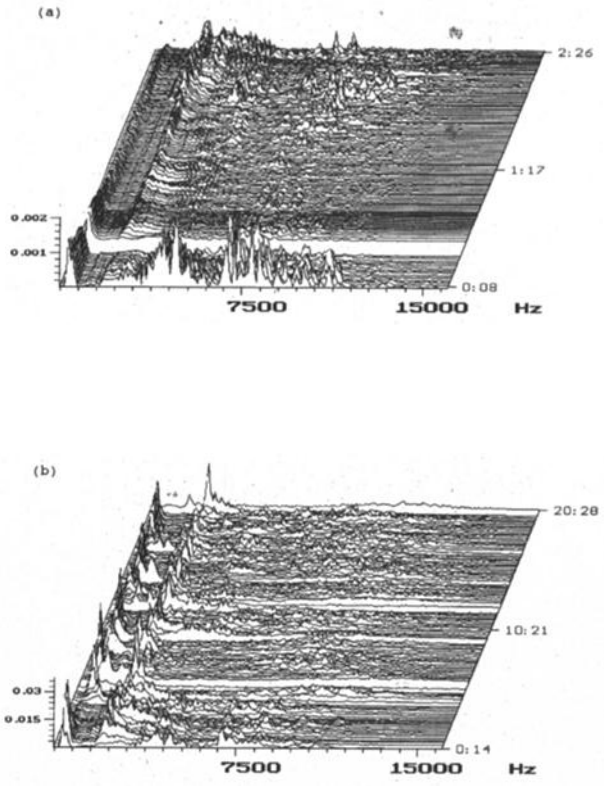

FIG. 6. (a) Waterfall display of sequential click spectra, showing power magnitude data over the first 120 clicks from bull male click sequence SW1. Ordinate rms power in $V^2\ s^{-1}$, abscissa frequency in hertz, and cascade axis time in minutes since the first click of the sequence. Each individual graph in the cascade in the power spectrum of a single click. (b) Waterfall display of normalized and averaged click spectra, showing overall power spectral distribution over the full duration of male click sequence SW1. Each individual graph in the cascade is the average of ten sequential normalized power spectra. Ordinate shows normalized power units; other axes as in (a).

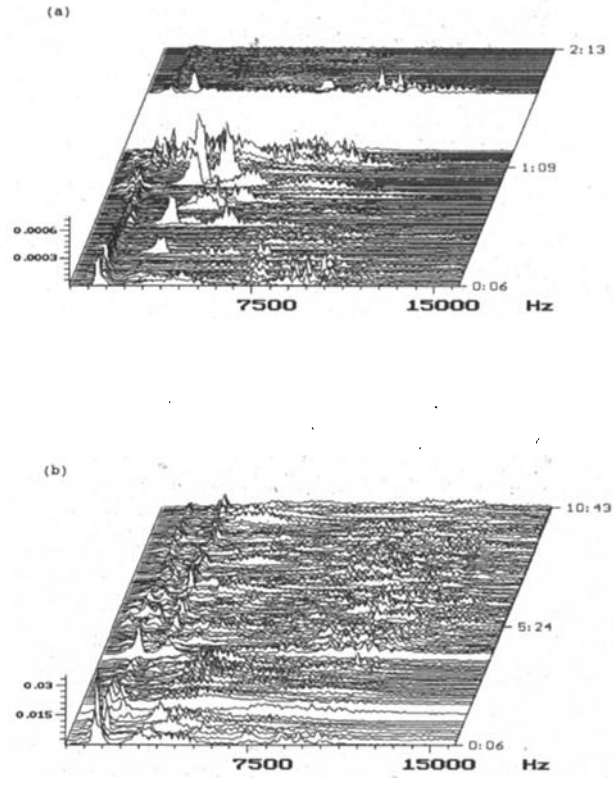

FIG. 7. (a) Waterfall display of sequential click spectra, showing power magnitude data over the first 120 clicks from female click sequence SW6; axes units are as in Fig. 6(a). (b) Waterfall display of normalized and averaged click spectra, showing overall power spectral distribution over the full duration of female click sequences SW6; axes units are as in Fig. 6(b).

0.0048) and for females is 1.9455 s^{-1} (standard error 0.0075).

B. Frequency domain analysis

Figure 6(a) shows a waterfall spectral plot for the first 120 clicks of the large bull male SW1, displayed as cascading time series of power magnitude spectra. Each graph in the cascade is the power magnitude spectrum of a single click, with time increasing from front to back along the cascade axis. Although all frequencies in the range 0–24 kHz were captured, only those components in the 0 to 16 kHz range are displayed, as virtually all the power is represented in this band. The plot reveals prominent peaks at approximately 400 Hz and 2 kHz; these frequencies have a tendency to dominate the spectrum during the early stages of the dive, although in Fig. 6(a) the first few clicks can be seen to contain substantial amounts of high-frequency energy. Note also that the power level of the low-frequency peak (400 Hz) tends to decay with time, whereas the 2 kHz peak does not follow a similar pattern. Energies are present throughout the higher frequency range, but are not concentrated at any sharply defined frequency. This sequence gives a clear illustration of early click characteristics for bull male sperm whales as recorded by surface hydrophones, since the original recordings were of very high signal-to-noise ratio. However, it should be remembered that the signals were recorded from behind diving animals and off the center of the main transmission beam, so are likely to be distorted with reference to on axis signals. Figure 6(b) shows the normalized and averaged power spectra of all clicks in sequence SW1, giving an overview of the spectral distribution over extended time periods. It can be seen that the 400 Hz and 2 kHz peaks are stable features throughout the duration of the sequence. Higher frequency components tend to occur in the 6 to 12 kHz band, with low levels above about 12 kHz. Figure 7(a) shows sequential power spectra of the clicks from the early stages of female click sequence SW6. As with the males, there is a concentration of low-frequency energy, but now seemingly restricted to a peak at about 1.2 kHz. The corresponding normalized and averaged spectra [Fig. 7(b)] reveals more clearly the trends over extended time periods. There is a tendency for energy to concentrate at 1.2 and 3 kHz, but the spectral peaks seem less sharply defined and less stable with time than the prominent peaks in the male spectrum. Higher frequency components are broadband and extend up to 15 kHz.

C. Time domain analysis

Figure 8 shows the results of filtering the click of a large bull male sperm whale in several incremental frequency

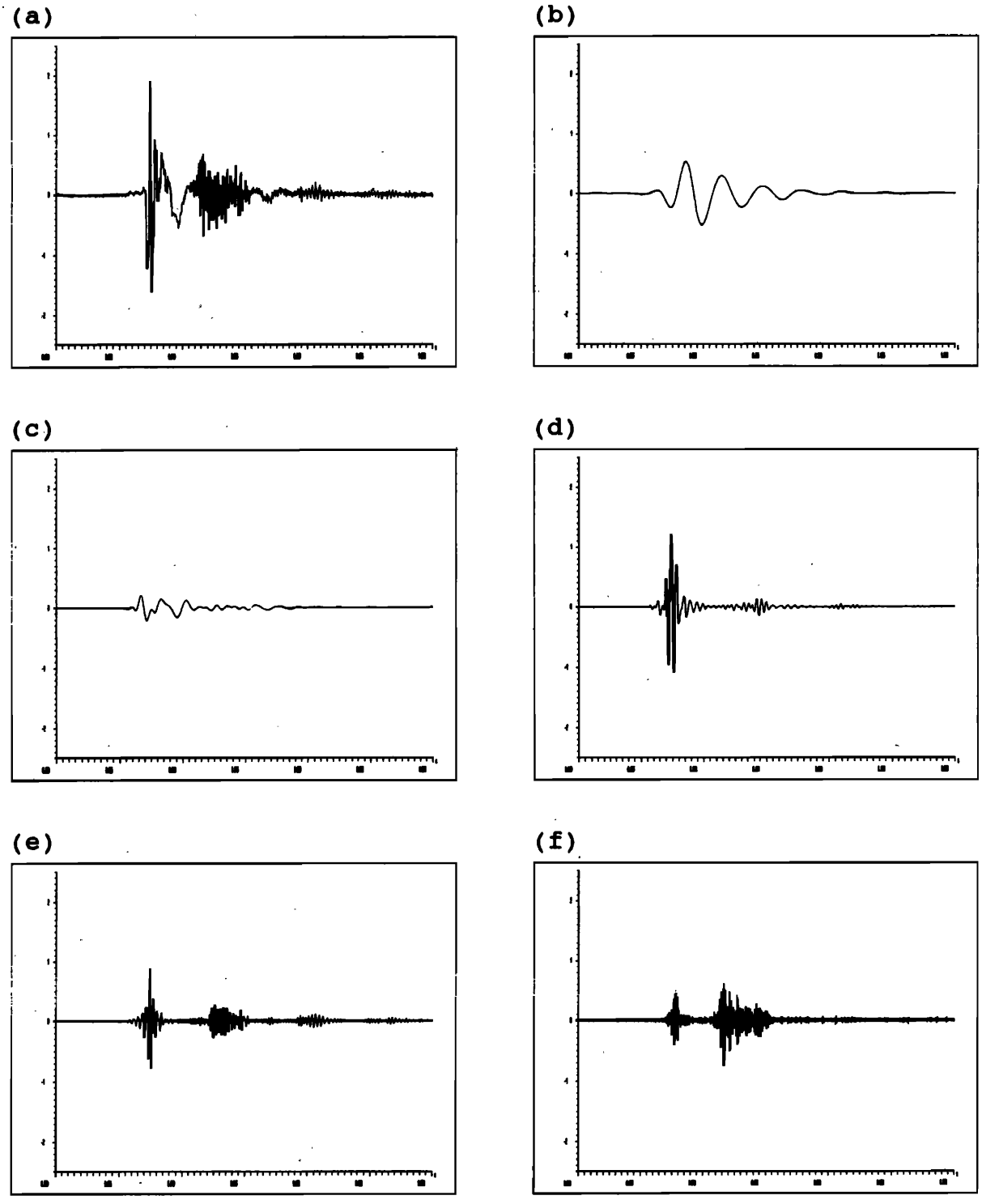

FIG. 8. Isolation of incremental frequency components within a bull male sperm whale click from the early stages of sequence SW2. The unfiltered signal is shown in (a) with subsequent plots showing isolation of components in the following bands: (b) 0–500 Hz, (c) 500 Hz–1.5 kHz, (d) 1.5–3 kHz, (e) 3–6 kHz, and (f) 6+ kHz. Time base is 30 ms in each plot; waveform units in volts.

bands. Figure 8(a) shows the unfiltered signal and (b) shows isolation of the low-frequency component (approx 400 Hz) with a 500 Hz low-pass FIR filter. The time domain form of this component is a heavily damped sinewave, and represents the typical form of the subkilohertz components of clicks from large males. There is no evidence of a repeating series of pulses, which is to be expected as the wavelength at this frequency, approximately 2.8 m in spermaceti oil ($c_{\text{spermaceti}} = 1400\ \text{m s}^{-1}$), is much larger than the dimensions of the proposed reflective surfaces of the frontal and distal air sacs (less than 1 m). The spectral plots in Fig. 6 indicate very little energy in the band between 500 Hz and 1.5 kHz, and Fig. 8(c) shows the fairly formless, low level components in this frequency band.

The second prominent peak in the male click spectra occurs at approximately 2 kHz, and encompasses a slightly broader range of frequencies, from about 1.5 to 3 kHz. These components were isolated with a bandpass FIR filter centered on 2.25 kHz with −3 dB points at 1.5 and 3 kHz. Figure 8(d) shows a typical result of isolating this narrow frequency band. The 2-kHz spike constitutes the sharp onset of the click, and energy at this frequency is largely confined to this spike of some 2 to 3 ms duration at the beginning of the click. Note also that a decaying pulse train is becoming visible, the wavelength of sound now some 0.7 m in spermaceti oil, at which reflection of sound from the frontal and distal air sacs is more feasible. The decay in energy with successive pulses is very rapid, as might be expected, and the pulse spacing is very regular, approximately 7.2 ms, which conforms with Norris and Harvey's hypothesis for an animal of this size (approximately 18 m). However, a clearly defined pulsed structure in this frequency band did not persist as dives progressed, although it would sharpen and degenerate intermittently.

Above 2 kHz there are no particularly prominent peaks in the click spectrum, so components at higher frequencies were inspected by progressively widening the filter bandwidths, namely, 3–6 kHz and finally 6+ kHz. A bandpass FIR filter centered at 4.5 kHz, with −3 dB points at 3 and 6 kHz, and a high-pass FIR filter with −3 dB points at 6 kHz were used, respectively. In the 3 to 6 kHz band the pulse structure is quite well defined and persists throughout the dive (e). However, it is notable that the envelope of the first and second pulses in the click become dissimilar, with the first pulse envelope retaining a short, spike form of some 2 to 3 ms duration, and the second pulse envelope becoming more spread across some 4–5 ms. In addition, it can be seen that the energy levels in successive pulses decay less rapidly than in the 1.5 to 3 kHz band. Inspection of the components above 6 kHz revealed a marked dissimilarity between the envelopes of first and second pulses in the click (f). The first pulse envelope is still confined to some 2 to 3 ms, whereas the second pulse now spreads across some 5–6 ms. There is also a large amount of energy in the second pulse, relative to the first, but multiples following this second pulse are extremely low in energy.

Figure 9 shows the result of applying identical filters to the click of a smaller female sperm whale. It can be seen (b) that components below 500 Hz are virtually absent, whereas in the male click a significant sinewave component can be isolated. The approximate equivalent to this sinewave is to be found in the 500 Hz to 1.5 kHz band (c), representing the low-frequency component of 1.2 kHz seen in the spectral plots (Fig. 7). Components from 1.5 to 3 kHz (d) show no evidence of a pulsed structure, unlike the male click. This might be expected as the dimensions of a mature female sperm whale head are considerably less than those of a bull male, and as such the reflective air sacs will also be proportionately smaller. A decaying pulsed structure emerges in the 3 to 6 kHz band (e), corresponding to wavelengths of 0.5–0.25 m. Components above 6 kHz (f) show, similar to the male click, a marked dissimilarity of envelope between first and second pulses. The clicks portrayed in these figures, while merely snapshots of the whole, show features that are generally representative of the vast majority of male and female sperm whale clicks sampled during this study.

D. Time–frequency analysis

Figure 10 shows the result of an STFT performed on one of the first clicks from the bull male sequence SW2. The picture confirms the complex time–frequency structure already visualized by digital filtering. The onset of the click shows energy at 400 Hz and 2 kHz, corresponding to the components isolated in Fig. 8. The underlying 400 Hz sinewave can be seen to reach peak energy about 2 ms after the 2 kHz component. The 2 kHz component decays very rapidly; the 400 Hz component persists for several milliseconds. Frequencies from 3 to 12 kHz can be seen to dominate some 6 ms after the initial click onset, with a peak at about 4 kHz. These high-frequency components have energy levels greater than occurred at the click onset in the corresponding bands.

III. DISCUSSION

It is assumed that the regular clicks of diving sperm whales are produced for the purposes of echolocation. The transient nature of the signals are typical of cetacea that are known to echolocate, i.e., dolphins, and the high-frequency components of the clicks should permit reasonable "imaging resolution," certainly discrimination of features less than half a meter across. The click rate graphs (Figs. 4 and 5) indicate that a fairly constant rate of clicking is maintained for the majority of the dive. This would tend to suggest that increased clicking rate with decreasing target range is unnecessary. Such an assumption is reasonable at long ranges of hundreds or thousands of meters, as target information will change very little from one echo-return to the next. At close ranges, however, the angular position of a target, especially escaping prey, will change more rapidly, and hence information on its position must be gained at an increasing rate (Suga, 1990). Increased click rates preceding creaks and the creaks themselves may be indicative of animals closing on prey.

Figures 4 and 5 show that the only overall increase in click rate occurred at the beginning of each dive, where click rates increase over a period of some 3 to 4 min. Initial click rates for both males and females are in the region of 0.5–1 s^{-1}, implying maximum sonaring ranges of 1500–750 m (c_{water}=1500 m s^{-1}) assuming that the whale emits a click upon immediate reception of the echo from its previous click. These ranges are comparable to the water depths in which the animals were feeding, and therefore the seabed proximity may have had some influence on the increase in click rate during the early dive stages. Sperm whales off the Azores dive almost vertically toward the seabed during these first few minutes and typical descent rates were measured by an echo sounder at approximately 100 m/min. Thus, in 4 min, a descending whale is 400 m closer to the seabed, shortening the total transmission path of a signal in the vertical plane by 800 m and thus the round trip time for a signal echo from the seabed by about 0.5 s.

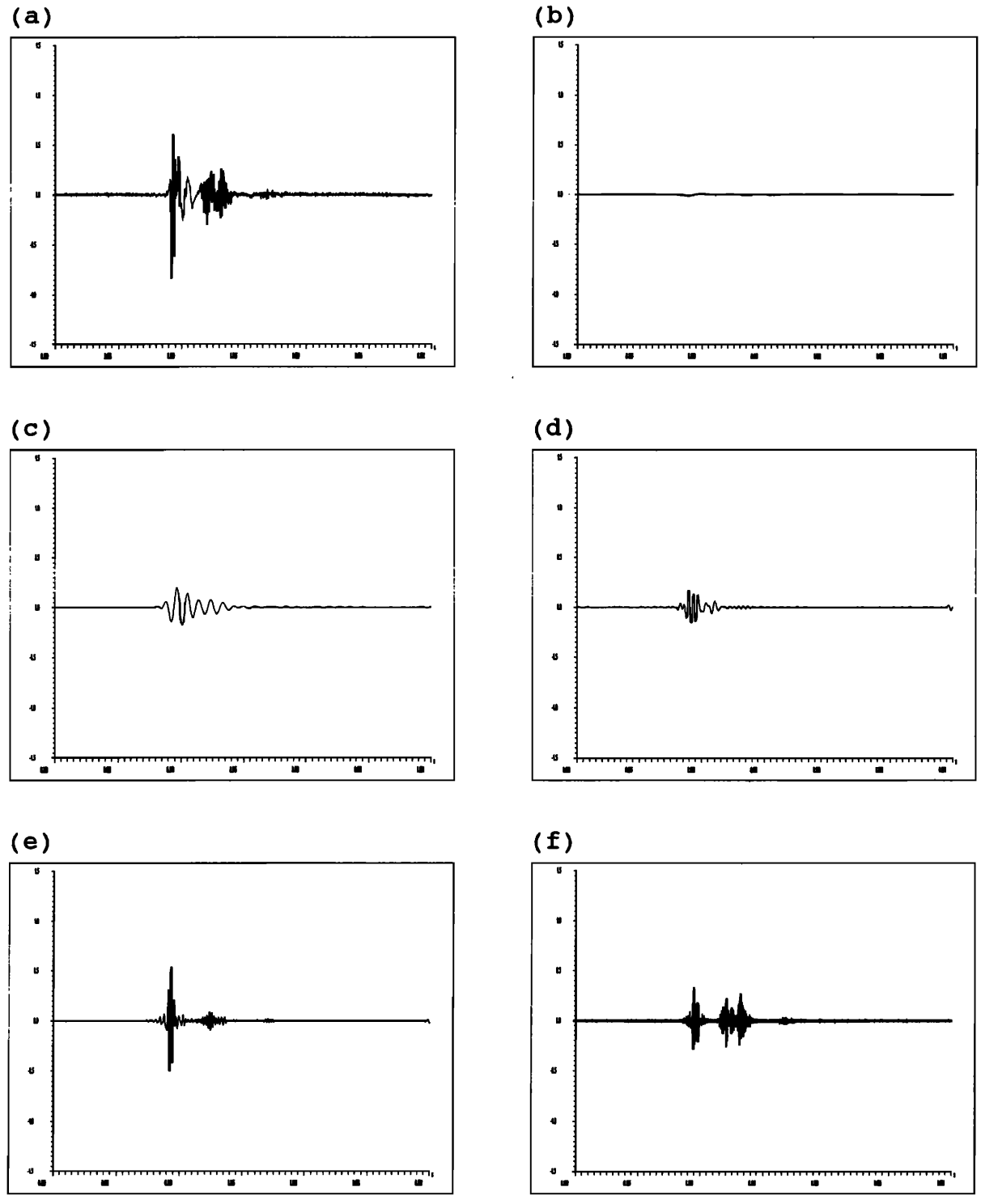

FIG. 9. Isolation of incremental frequency components within a female sperm whale click from the early stage of sequence SW6. Filter bands, time bases, and waveform units are as in Fig. 8.

It is likely that a multiplicity of factors is involved, although the depths and proposed transmission ranges involved certainly bear a feasibly close relationship. Using the click rate data from 4 min onward, overall mean click rates have been calculated for bull males at 1.1713 s^{-1} and for females at 1.9455 s^{-1}. These equilibrium rates imply maximum sonaring ranges of about 640 m for males and about 385 m for females. The reason for this difference is not entirely clear. One possible explanation is that, being physically much larger than females, bull males can produce more powerful clicks, and hence gain "useful" target information over greater ranges than females. Greater range would, naturally, imply that the interval between clicks would increase to await the more distant echo returns. Since the male head is

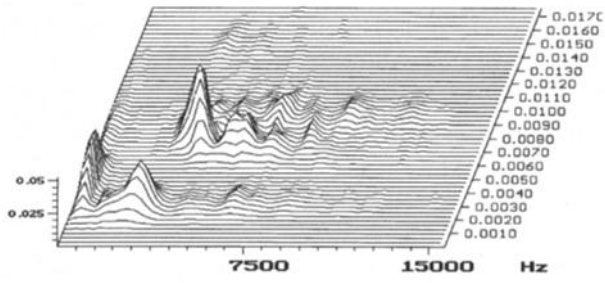

FIG. 10. Waterfall plot showing short-time Fourier-transform sequence across the duration of a click from the bull male sperm whale of sequence SW2. Ordinate shows rms power in $V^2 s^{-1}$, abscissa shows frequency in hertz, and the cascade axis shows the time shift of the FFT window in seconds from a point immediately preceding the click onset.

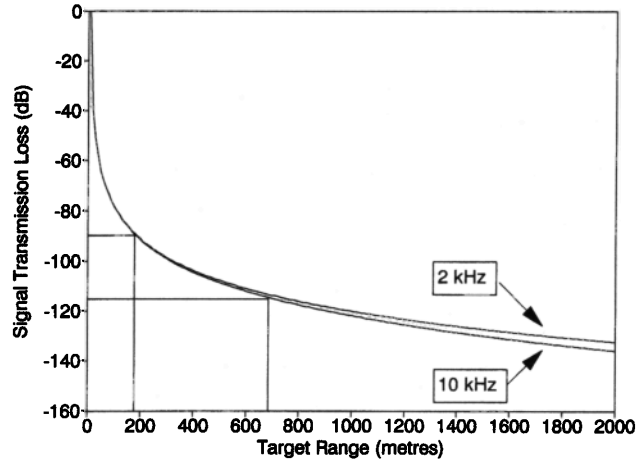

FIG. 11. Theoretical curves for total transmission loss due to spherical spreading and absorption of 2 and 10 kHz signals in deep oceanic water. Estimated target detection ranges for male sperm whales, from application of a noise limited form of the sonar equation, are also indicated.

physically much more massive, it may allow the buildup and discharge of greater volumes/pressures of air during sound production, resulting in higher source level emissions.

The detection threshold of the sperm whale sonar system will depend on a number of factors, which can be summarized in a noise limited form of the sonar equation, quoted after Au (1993). Thus

$$DT = SL - 2TL + TS - (NL - DI). \quad (1)$$

All parameters are expressed in dB, the detection threshold (DT) being the difference between the received echo level and the received noise spectral density at which the animal is just able to distinguish the echo. Assume in the following example that DT=0, i.e., the whale can just discern an echo when the echo level is equal to the ambient noise level.

The click source level can be estimated from the data of Levenson (1974) who quotes a mean broadband source level for sperm whale clicks of 171.2 dB *re*: 1 μPa, although it is not stated whether this refers to peak or rms values. The size, sex, and maturity of the target whales in Levenson's study are also unclear, although the location would suggest males and the fact that four animals were grouped would suggest juvenile males. Large bulls tend to be solitary but the younger males seem to have some cohesion in small groups away from breeding latitudes. In addition, Levensons' paper shows a couple of oscilloscope traces of the clicks from target animals which seem to show an IPI of 6 to 7 ms, which would correspond to fairly large males. Assume, therefore, a male click source level (SL) of 170 dB at a given frequency.

The transmission loss (TL) depends on the range, directional characteristics, and frequency of the sound. Absorption losses in seawater for the frequencies under consideration are very small, typically between 10^{-4} and 10^{-3} dB per meter. For free-field spherical spreading, which will be assumed here, the two-way transmission loss (two ways as an echo return is required from the target) follows a $-40\log(r)$ curve. Figure 11 shows theoretical curves for two-way transmission loss due to the combined effects of spherical spreading and absorption in deep oceanic water for 2 and 10 kHz signals.

Target strength can be estimated from published values for other species similar to the expected prey of the sperm whale. Prey species are primarily deep sea squid of which little is known other than the beaks retrieved from the stomachs of predators (including sperm whales). Smith (1954) measured target strength for individual squid, *Loligo pealei*, of -40 dB at 10 kHz. As this value is for a single squid, -40 dB will be assumed as the minimum target strength for sperm whale prey.

Ambient noise levels in the deep sea in the three frequency decades between 100 Hz and 100 kHz are primarily determined by surface agitation (Coates, 1990). Ambient noise levels in the low kilohertz region are generally between 50 and 70 dB *re*: 1 μPa depending on surface conditions. At 10 kHz these noise levels would be expected to drop by about 15 dB. As the recordings used in this study were made during periods of good weather, ambient noise levels of 50 and 35 dB will be assumed at 2 and 10 kHz, respectively. These frequencies are of interest as a strong component of the male click centers around 2 kHz and there are significant higher frequency components around the 10 kHz region. As the attenuation of surface agitation noise will depend only on absorption losses, which are very small at these frequencies, the effects can be neglected and the sound field considered isotropic throughout, although in practice scattering layers may significantly influence the noise spectral density at depth.

Finally, the receiving beam directivity index (DI) must be estimated. If the receiving system of the sperm whale has directional properties, which seems likely considering the case for small cetaceans, then it will receive less noise in an isotropic noise field than will an omnidirectional receiver. Unfortunately, measuring the receiving beam of a deep diving sperm whale presents obvious problems and hence there are no such data available. The directivity index will therefore be estimated using the data for other species, namely, dolphin. As the gross morphology of the sperm whale skull and jaw bones is analogous to those of dolphins, similar directional hearing characteristics may also pertain. Such an assumption could be valid; the wavelengths of sound used by sperm whales are an order of magnitude longer than those corresponding to the peak frequencies at which dolphin so-

nars operate, but the physical dimensions of the sperm whale head are an order of magnitude greater to compensate. Au (1993) quotes receiving beam DIs of 10.4, 15.3, and 20.6 dB for *Tursiops truncatus* at 30, 60, and 120 kHz, respectively. Therefore, a crude assumption will be made that sperm whales have receive beam DIs of 10 and 20 dB at 2 and 10 kHz, respectively.

For 10 kHz signal components Eq. (1) may now be followed as

$$DT=SL-2TL+TS-(NL-DI),$$

$$0=170-2TL-40-(35-20),$$

$$0=115-2TL,$$

$$2TL=-115.$$

Inspection of Fig. 11 indicates that the target range corresponding to a two-way transmission loss of −115 dB at 10 kHz is about 680 m, just over the maximum implied for males by click rates in this study. One additional feature of dolphin sonar is that there always exists a lag time between the two-way transit time for a signal-echo return from a target and the actual time interval between echolocation clicks. Au (1993) quotes these lag times, which remain virtually constant with target range, between 19 and 45 ms for *Tursiops*. If similar time lags occur in sperm whale sonar, they would reduce the implied maximum sonaring ranges for male sperm whales by about 14 and 34 m, respectively, thereby pushing the sonar equation estimate some 50–80 m in excess of the theoretical maximum. Even so the estimate is fairly close considering the number of assumptions made. If the calculation through Eq. (1) is repeated for 2-kHz signal components of SL=170 dB, with NL=50 dB and DI=10 dB, the range performance drops to just under 200 m. This result, with the obvious target resolution advantages associated with higher frequencies, implies that the components of sperm whale clicks around 10 kHz are of greater utility in echolocation than those around 2 kHz.

The assumption of spherical spreading is probably justified in the low kilohertz region considering the wavelength of sound at these frequencies. At 10 kHz the transmission beam may become more focused, which would have the effect of reducing TL and making the sonar range estimate even greater. However, ambient noise is probably underestimated here as there would almost certainly be additional reflection and reverberation noise from scattering layers and other discontinuities. Indeed the sperm whale sonar system may be reverberation limited in this sense. The detection threshold and receive beam directivity index are also a matter of some speculation, so there is limited value in substituting values *ad infinitum* until the number of unknown parameters is reduced by future experimental results.

The previously reported existence of a multiple pulsed structure within sperm whale clicks has been confirmed by this work. Filtering often shows the second pulse envelope of high-frequency components to be excessively large compared to that expected from a simple reflection. In fact, reflection is far from simple in most practical situations, especially in a system as complex as the sperm whale head. At frequencies upward of 3 kHz the transmit beam of male sperm whales may start to exhibit directional properties. For a diving sperm whale a component of this beam would reach surface hydrophones directly, while a component will be reflected forward from the air filled nasofrontal sac. A component of this reflected pulse will again undergo reflection, this time directed backward from the distal air sac at the front of the head. This component should, according to the Norris and Harvey hypothesis, form the second pulse in the click received at the surface. The rearward radiation pattern of the initial sound pulse and that of the first reflection from the distal sac are likely to be dissimilar. If the initial signal had a major rearward transmission lobe along the axis of the whale, and the animal were oriented such that the surface hydrophones were outside this major lobe, then one might expect to receive a relatively small amount of energy from the initial pulse. Further, if the radiation pattern of the first reflection from the distal sac were oriented more on axis toward the hydrophone, then perhaps one might expect to receive more energy. The lengthening of the second pulse envelope is also indicative of a complex reflection.

If such a mechanism were indeed taking place, then the relative envelopes and energy contents of the first and second pulses in the click received by surface hydrophones might provide cues as to the orientation of submerged whales. The third and subsequent pulses in the click should decay rapidly with respect to the second, as observed, for they will be the product of passive reflection with similar directional characteristics to the second pulse, whereas the first pulse is actively generated and may have a different radiation pattern. Such a view is rather speculative at present but the idea would be a good subject for computer modeling to estimate transmission beams for the various components of sperm whale clicks.

It is perhaps also possible that components of the initial pulse, reflected forward from the nasofrontal air sac, act to resonate the distal air cavity, thus facilitating synchronous vibration of the museau de singe lips and resulting in a second burst of high-frequency energy. Mackay (1980) made the similar suggestion that a returning sound pulse would resonate the air in the distal sac and cause the museau de singe lips to vibrate at a preferred frequency, as do the lips of a trumpet player. In addition, Mackay suggested that the spermaceti organ may behave like an organ pipe, with quarter- or half-wavelength resonance. This suggestion seems less likely as the spermaceti organ is not isolated from its surroundings by discontinuities of significantly different acoustic impedance. The air spaces of the frontal and distal air sacs, and even the right nasal passage, are much stronger candidates for resonance effects. The precise form and dimensions of these air cavities in the living, vocalizing animal are a matter of some conjecture, which makes estimates of resonant frequencies difficult. It would seem, though, that increased energy output at high frequencies may well be beneficial to the animals' sonar performance, unless the performance is reverberation limited.

The function, if any, of the multiple pulsed structure of sperm whale clicks is unclear. Norris and Harvey (1972) suggested that several pulses produce several echo returns which

act to cause "summation" in the auditory processing in the sperm whale brain, making an echo sound "louder" to the animal. They likened this effect to firing a gun in an enclosed space, where the multiple reverberant echoes from the walls make the report sound far louder than when the same gun was fired in an open space. It is questionable whether sufficient reverberation is present in a distant echo return to produce a similar effect in the sperm whale auditory system. A quite different idea (Gordon, personal communication) suggests that the pulsed structure serves an intraspecific social function. As the pulse spacing within a click is a function of the head length, it is suggested that the click pulse spacing may indicate to other whales the size, and hence dominance, of the animal. For instance, conveying such length information may facilitate competition among males without the need for physical combat.

The observed oscillations in click rate are also somewhat puzzling, and defy immediate explanation. It is interesting to note that Au (1980) found a similar pattern in the click rates of bottlenose dolphins echolocating in open waters, although no explanation was offered for this phenomenon. It may be that a whale continually redirects its receive beam while searching for food, thereby scanning regions of differing noise level which dictate maximum detection ranges. It may also be that the oscillations are random and have no significance whatsoever; a whale may not be able, or wish, to time its clicks more precisely for long-range echolocation.

IV. CONCLUSIONS

Sperm whale clicks are complex, broadband transients. Both male and female sperm whales have stable, or quasistable, peaks in their click spectra which seem to be associated with the animal's size. Individual clicks contain multiples of the initial pulse above a certain threshold frequency at which the wavelength of sound approximates to the dimensions of the frontal and distal air sacs. In addition, the spacing of these multiples is in accordance with the expected longitudinal separation of the air sacs within the head of the animal, using published and original values of sound velocity in spermaceti oil. As such the outline model of sound production proposed by Norris and Harvey (1972) is supported. However, the envelope and energy content of the multiples changes with increasing frequency, suggesting some form of off axis distortion or a secondary active mechanism of high-frequency sound production or reinforcement. It is also suggested that these higher frequency components are of greater utility in echolocation than the low-frequency components of the clicks.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Jonathan Gordon and the International Fund for Animal Welfare for making available the acoustic data used in this study, and also for their assistance in providing fieldwork opportunities for the senior author. Thanks are also expressed to Dr. Jim Bennell of the School of Ocean Sciences for much useful discussion and technical assistance.

Alder-Fenchel, H. S. (**1980**). "Acoustically derived estimate of the size distribution for a sample of sperm whales, *Physeter catodon*, in the Western North Atlantic," Can. J. Fish. Aquat. Sci. **37**, 2358–2361.

Au, W. W. L. (**1980**). "Echolocation signals of the Atlantic Bottlenose Dolphin, *Tursiops truncatus*, in open waters," in *Animals Sonar Systems*, edited by R. G. Bunsel and J. F. Fish (Plenum, New York), pp. 251–282.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer-Verlag, New York).

Backus, R. H., and Schevill, W. E. (**1966**). "Physeter clicks," in *Whales, Dolphins and Porpoises*, edited by K. S. Norris (University of California, Berkeley, CA), pp. 510–528.

Bunsel, R. G., and Dziedzic, A. (**1967**). "Observations sur le comportment et les emissions acoustiques du cachalot lors de la chasse," Bocagiana **14**, 1–15.

Clarke, M. R. (**1979**). "The head of the sperm whale," Sci. Am. **240**(1), 106–117.

Clarke, M. R., Martins, H. R., and Pascoe, P. (**1993**). "The diet of sperm whales, *Physeter macrocephalus Linnaeus 1758*, off the Azores," Philos. Trans. R. Soc. London Ser. B **339**, 67–82.

Coates, R. F. W. (**1990**). *Underwater Acoustic Systems*, MacMillan New Electronics Introductions to Advanced Topics (MacMillan, London).

Gordon, J. C. D. (**1987**). "The behaviour and ecology of sperm whales off Sri Lanka," Ph.D. thesis, University of Cambridge.

Gordon, J. C. D. (**1991**). "Evaluation of a method for determining the length of sperm whales, *Physeter catodon*, from their vocalisations," J. Zool. London **224**, 301–314.

Levenson, C. (**1974**). "Source level and bistatic target strength of the sperm whale, *Physeter catodon*, measured from an oceanographic aircraft," J. Acoust. Soc. Am. **55**, 1100–1103.

Lockyer, C. (**1977**). "Observations on diving behaviour of the sperm whale, *Physeter catodon*," in *A Voyage of Discovery*, edited by M. Angel (Pergamon, Oxford), pp. 591–609.

Lynn, P. A., and Fuerst, W. (**1992**). *Introductory Digital Signal Processing, With Computer Applications* (Wiley, New York).

Mackay, R. S. (**1980**). "A theory of the spermaceti organ in sperm whale sound production," in *Animal Sonar Systems*, edited by R. G. Bunsel and J. F. Fish (Plenum, New York), pp. 937–940.

Mohl, B., Larsen, E., and Amundin, M. (**1981**). "Sperm whale size determination: Outlines of an acoustic approach," FAO Fish. Ser. **5**, 327–332.

Mohl, B., and Amundin, M. (**1991**). "Sound production in odontocetes with emphasis on the harbour porpoise, *Phocoena phocoena*," Mats Amundin Doctoral dissertation, Department Zoology, Stockholm University.

Norris, K. S., and Harvey, G. W. (**1972**). "A theory for the function of the spermaceti organ of the sperm whale, *Physter catodon*," in *Animal Orientation and Navigation*, edited by S. R. Galler, K. Schmidt-Koenig, G. J. Jacobs, and R. E. Belleville, NASA special Publications No. 262, Washington, DC, pp. 397–417.

Smith, P. F. (**1954**). "Further measurements of the sound scattering properties of several marine organisms," Deep-Sea Res. **2**, 71–79.

Suga, N. (**1990**). "Biosonar and neural computation in bats," Sci. Am. **262**(b), 34–41.

Watkins, W. A. (**1977**). "Acoustic behaviours of sperm whales," Oceanus **20**, 50–58.

Watkins, W. A., Daher, M. A., Fristrup, K. M., Howald, T. J., and Disciara, G. N. (**1993**). "Sperm whales tagged with transponders and tracked underwater by sonar," Mar. Mammal Sci. **9**, 55–67.

Weilgart, L. S. (**1990**). "Vocalisations of the sperm whale, *Physeter macrocephalus*, off the Galapagos Islands as related to behavioural and circumstantial variables," Ph.D. thesis, Dalhousie University Halifax, Nova Scotia.

Weilgart, L. S., and Whitehead, H. (**1988**). "Distinctive vocalisations from mature male sperm whales, *Physeter macrocephalus*," Can. J. Zool. **66**, 1931–1937.

Whitehead, H., and Weilgart, L. S. (**1990**). "Click rates from sperm whales," J. Acoust. Soc. Am. **87**, 1798–1806.

Worthington, L. V., and Schevill, W. E. (**1957**). "Underwater sounds heard from sperm whales," Nature (London) **180**, 291.